Simulations and a conditional limit theorem for intermediately subcritical branching processes in random environment

Ch. Böinghoff *   G. Kersting *

September 7, 2012

Abstract

Intermediately subcritical branching processes in random environment are at the borderline between two subcritical regimes and exhibit a particularly rich behavior. In this paper, we prove a functional limit theorem for these processes. It is discussed together with two other recently proved limit theorems for the intermediately subcritical case and illustrated by several computer simulations.

MSC 2000 subject classifications. Primary 60J80, Secondary 60F17

Key words and phrases. Branching process, random environment, functional limit theorem

Running head. Branching processes in random environment

*Fachbereich Mathematik, Universität Frankfurt, Fach 187, D-60054 Frankfurt am Main, Germany , boeinghoff@math.uni-frankfurt.de, kersting@math.uni-frankfurt.de

1 Introduction and main results

Branching processes in random environment are a discrete time model for the development of a population of individuals. In contrast to Galton-Watson processes, it is assumed that individuals are exposed to a random environment, which influences the reproductive success of the individuals and varies from one generation to the next in an i.i.d. manner. Given the environment, all individuals reproduce independently according to the same mechanism.

More precisely, let Q be a random variable taking values in the space Δ of probability measures on $\mathbb{N}_0$. Equipped with the total variation metric, Δ is a Polish space. An infinite sequence $\Pi = (Q_1, Q_2, \ldots)$ of i.i.d. copies of Q is called *random environment*. We denote by Q_n the offspring distribution of an individual in generation $n-1$ and by Z_n the number of individuals in generation n. A sequence of $\mathbb{N}_0$-valued random variables $Z_0, Z_1, \ldots$ is called a *branching process in the random environment (BPRE)* Π, if Z_0 is independent of Π and given Π the process $Z = (Z_0, Z_1, \ldots)$ is a Markov chain with

$$\mathcal{L}\big(Z_n \mid Z_{n-1} = z, \Pi = (q_1, q_2, \ldots)\big) \;=\; q_n^{*z} \tag{1.1}$$

for every $n \in \mathbb{N}$, $z \in \mathbb{N}_0$ and $q_1, q_2, \ldots \in \Delta$, where q^{*z} is the z-fold convolution of the measure q. The corresponding probability measure on the underlying probability space will be denoted by $\mathbb{P}$. In the following we assume that the process starts with a single founding ancestor, $Z_0 = 1$ a.s. Throughout the paper, we shorten $Q(\{y\}), q(\{y\})$ to $Q(y), q(y)$.

As it turns out, the fine structure of the offspring distributions is of secondary importance for the asymptotics of the BPRE. More precisely, the asymptotics of Z is mainly determined by the *associated random walk*, which only contains information on the mean offspring number in each generation. The associated random walk $S = (S_n)_{n\geq 0}$ is the random walk having initial state $S_0 = 0$ and increments $X_n = S_n - S_{n-1}$, $n \geq 1$ defined by

$$X_n \;=\; \log m(Q_n),$$

where

$$m(q) \;=\; \sum_{y=0}^{\infty} y q(y)$$

is the mean of the offspring distribution $q \in \Delta$. Assuming $Z_0 = 1$ a.s., it results that the conditional expectation of Z_n given the environment Π can be written

as

$$\mathbb{E}[Z_n \,|\, \Pi\,] \;=\; \prod_{k=1}^{n} m(Q_k) \;=\; e^{S_n} \quad \mathbb{P}\text{–a.s.} \tag{1.2}$$

Averaging over the environment yields

$$\mathbb{E}[Z_n] \;=\; \mathbb{E}[e^{S_n}] \;=\; \mathbb{E}[e^X]^n \tag{1.3}$$

with $X = \log m(Q)$.

There are several phase transitions present in the class of BPRE. They already become visible, when looking at the asymptotic survival probability. From Jensen's inequality for $0 \le \lambda \le 1$

$$\mathbb{P}(Z_n > 0) \le \mathbb{E}[Z_n^\lambda] = \mathbb{E}\big[\mathbb{E}[Z_n^\lambda \mid \Pi]\big] \le \mathbb{E}\big[\mathbb{E}[Z_n \mid \Pi]^\lambda\big] = \mathbb{E}[e^{\lambda X}]^n \ .$$

Indeed under quite general conditions (see [1, 10, 13, 12]) it holds that

$$\lim_{n\to\infty} \mathbb{P}(Z_n > 0)^{\frac{1}{n}} = \inf_{0\le\lambda\le1} \mathbb{E}[e^{\lambda X}] \ .$$

The formula suggests where the phase transitions are located. This depends on the value $\lambda_{\min}$, where the moment generating function $\lambda \mapsto \mathbb{E}[e^{\lambda X}]$ has its minimum. One phase transition appears, when $\lambda_{\min} = 0$, which essentially amounts to the condition $\mathbb{E}[X] = 0$. Then S is a recurrent random walk, and Z is called a *critical* BPRE. For a detailed study we refer to [5].

The other phase transition turns up when $\lambda_{\min} = 1$, which matches to the condition

$$\mathbb{E}[Xe^X] = 0 \ .$$

This condition in turn implies $\mathbb{E}[e^X] < 1$ and $\mathbb{E}[X] < 0$. Then Z is called an *intermediately subcritical* BPRE and S is a transient random walk with negative drift.

In the other cases $\lambda_{\min} < 0$, $0 < \lambda_{\min} < 1$ and $\lambda_{\min} > 1$ the BPRE Z is called *supercritical*, *weakly subcritical* and *strongly subcritical*, respectively, a classification, which goes back to Afanasyev [1] and Dekking [10].

As one would expect, BPREs exhibit a particularly rich behavior in the two transition cases. Here we focus on the intermediately subcritical case, namely on the behavior of the process up to time n, conditioned on the event $\{Z_n > 0\}$, in the limit $n \to \infty$. A main concern of our paper is to exemplify its features by means of computer simulations. In doing so we shall discuss three functional

limit theorems, which underly these simulations. Two are taken from the publication [4] (being the main basis of the present paper). The other limit theorem is proved below, which makes the second part of the paper. Intermediately subcritical BPREs have also been studied in [2, 12, 18]. For a comparative discussion we refer the reader to [9].

For an intermediately subcritical BPRE it is natural to introduce a change to the probability measure $\mathbf{P}$, given by its expectation

$$\mathbf{E}[\varphi(Q_1,\ldots,Q_n,Z_0,\ldots,Z_n)] \;=\; \gamma^{-n}\mathbb{E}\big[\varphi(Q_1,\ldots,Q_n,Z_0,\ldots,Z_n)e^{S_n}\big]$$

for any $n \in \mathbb{N}$ and any measurable, bounded function $\varphi : \Delta^n \times \mathbb{N}_0^{n+1} \to \mathbb{R}$, with $\gamma^n = \mathbb{E}[e^{S_n}] = \mathbb{E}[Z_n]$, thus

$$\gamma \;=\; \mathbb{E}[e^X] \; .$$

The condition $\mathbb{E}[Xe^X] = 0$ translates to

$$\mathbf{E}[X] \;=\; 0 \; .$$

Therefore S becomes a recurrent random walk under $\mathbf{P}$.

Let us pass to the assumptions. For $a \in \mathbb{N}$ denote

$$\zeta(a) \;=\; \sum_{y=a}^{\infty} y^2 Q(y)\Big/ m(Q)^2 \; , \quad a \in \mathbb{N} \; ,$$

which we refer to as the standardized truncated second moment of Q.

Assumption A. *Let X be non-lattice with*

$$\mathbf{E}[X] \;=\; 0 \; , \; 0 < \mathbf{E}[X^2] < \infty \; .$$

Moreover let

$$\mathbf{E}\big[(\log^+ \zeta(a))^{2+\varepsilon}\big] \;<\; \infty$$

for some $a \in \mathbb{N}$ and $\varepsilon > 0$, where $\log^+ x = \log(x \vee 1)$.

The condition $\mathbf{E}[X^2] < \infty$ is taken for convenience here. In the second half of the paper we shall replace it by a weaker assumption. See [5] for examples where the last assumption is fulfilled. In particular our result holds for binary branching processes in random environment (where individuals have either two children or none) and for cases where Q is a.s. a Poisson distribution or a.s. a geometric distribution.

Our first functional limit theorem concerns the stochastic processes $S^n = (S^n_t)_{0\le t\le 1}$, $n \in \mathbb{N}$, given by

$$S^n_t = S_{nt}\ , \quad 0 \le t \le 1\ .$$

Here and in the sequel we always shorten $\lfloor nt \rfloor$ to nt. Donsker's theorem states that

$$\frac{S^n}{\sigma\sqrt{n}} \xrightarrow{d} W$$

in distribution on the Skorohod space $D[0,1]$, where $W = (W_t)_{0\le t\le 1}$ is a standard Brownian motion and

$$\sigma^2 = \mathbf{E}[X^2]\ .$$

We denote by

$$W^c = (W^c_t)_{0\le t\le 1}$$

a process, which we call here a *conditional Brownian motion*, that is a Brownian motion conditioned to take its minimal value at $t = 1$.

Theorem 1.1. *Under Assumption A, it holds that as $n \to \infty$,*

$$\Big(\frac{S^n}{\sigma\sqrt{n}}\ \Big|\ Z_n > 0\Big) \xrightarrow{d} W^c$$

in the Skorohod space.

This theorem is taken from [4][Theorem 1.3]. The statement turns out to be characteristic for intermediately subcritical BPREs. It describes the impact on the random environment resulting from conditioning on the event $\{Z_n > 0\}$. Since $\mathbb{P}(Z_n > 0) \le \mathbb{E}[Z_n] = \mathbb{E}[e^X]^n$, note that $\mathbb{P}(Z_n > 0)$ is exponentially small such that we are in the range of large deviations. As usual there are distinct scenarios leading to different exponential rates for the probabilities, i.e. they require different 'costs'. Let us discuss this trade-off in detail.

First it is to be expected that Z_n is asymptotically of order $O_P(1)$ conditioned on the event $\{Z_n > 0\}$, since it would be too costly to build up a larger population. This enforces that $S_0, \dots, S_n$ have their minimum close to the end, because otherwise the population would have the chance for a late growth. Theorem 1.1 confirmes this consideration, yet the same phenomenon turns up also for the weakly and strongly subcritical case, see [3, 6].

Next note that among random walk paths $S_0, \dots, S_n$ with a late minimum one can imagine two strategies for the BPRE Z to survive until time n. Either

S builds up one big upward excursion. This is difficult to realize for a random walk with negative drift, but it provides an environment in which the branching process easily survives. Or S is on and on decreasing. This is readily realized for a random walk with negative drift, but gives the branching process a hard time to survive. Now the first alternative is realized for weakly subcritical and the second for strongly subcritical BPREs, see [3, 6]. Theorem 1.1 indicates that in the intermediately subcritical case both possibilities compete with each other, which means that they are equally costly. Indeed upward excursions alternate with decreasing ladder points within W^c.

Theorem 1.1 is our basis for the computer simulations of the conditional BPRE, given $\{Z_n > 0\}$. Since this event has exponentially small probability, direct simulations are not realizable. Also an access via a suitable Doob h-transform seems out of reach. Therefore we present an approximate solution by first simulating a conditional random walk path $S_0^c, \dots, S_n^c$ that is a random walk path $S_0, \dots, S_n$ conditioned to have its minimum at time n. Here we also rely on additional information supplied by [4][Theorem 1.3]: If τ_n denotes the moment, when $S_0, \dots, S_n$ attains its minimum for the first time, then $n - \tau_n$ given $Z_n > 0$ is convergent in distribution for $n \to \infty$. Given $S_0^c, \dots, S_n^c$ we generate the random environment. We choose a situation where the random walk completely determines the environment (otherwise we could not simply rely on Theorem 1.1). From a computational point of view it is convenient to choose geometric offspring distributions. Then given the environment the branching process $1 = Z_0, Z_1, \dots, Z_n$ conditioned to survive is efficiently generated by a general construction of the conditioned BPRE due to Geiger [11]. Altogether we replace the annealed situation in a way by a related quenched setting, which admittedly is only a substitute. The details are presented in Section 2.

Remark. The asymptotic shape of the limit distribution of $n - \tau_n$ given the event $\{Z_n > 0\}$ can be easily derived: For $0 \le k \le n$

$$\begin{aligned}&\mathbb{P}(n - \tau_n = k \mid Z_n > 0)\\ &\quad \le \frac{\mathbb{P}(\tau_{n-k} = n-k, Z_{n-k} > 0)\mathbb{P}(\min(S_1, \dots, S_k) \ge 0)}{\mathbb{P}(Z_n > 0)} .\end{aligned}$$

From [4] $\mathbb{P}(Z_{n-k} > 0) \sim \gamma^k \mathbb{P}(Z_n > 0)$ and $\mathbb{P}(\tau_n = n \mid Z_n > 0)$ has a positive limit. From [3][Proposition 2.1]

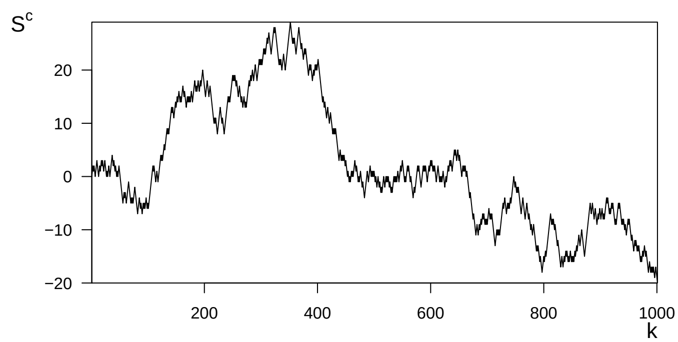
$$\mathbb{P}(\min(S_1, \dots, S_k) \ge 0) = \gamma^k \mathbf{E}[e^{-S_k}; \min(S_1, \dots, S_k) \ge 0] \sim c'\gamma^k k^{-3/2}$$

with some $c' > 0$. Thus there is a $c > 0$ such that for all $k \geq 0$

$$\lim_{n\to\infty} \mathbb{P}(n - \tau_n = k \mid Z_n > 0) \leq \frac{c}{k^{3/2}} \ .$$

With some effort this upper estimate can be refined to an asymptotic equality.

The following picture shows a path $S_0^c, \ldots, S_n^c$ of length $n = 1000$.



From Theorem 1.1 we expect the following behavior of $Z_0, Z_1, \ldots, Z_n$, conditioned on $Z_n > 0$: At decreasing ladder points of S the size Z_k is close (or equal) to 1, whereas during upward excursions of S the population size Z_k is tied to the up and down of its conditional expectation $\mathbb{E}[Z_k \mid \Pi] = e^{S_k}$ a.s. More precisely one would suspect that in either case $\log Z_k$ is close to $S_k - \min_{j\leq k} S_j$, which is the height of the random walk relative to its height at the last ladder point. This leads us to conjecture that as $n \to \infty$

$$\Big(\frac{1}{\sigma\sqrt{n}} \log Z_{nt} \ \Big| \ Z_n > 0 \Big) \xrightarrow{d} W_t^r$$

for $0 \leq t \leq 1$, where

$$W_t^r = W_t^c - \min_{s\leq t} W_s^c \ , \quad 0 \leq t \leq 1 \ ,$$

is the conditional Brownian motion *reflected* at its current minimum. Now a famous result of Lévy says that for the unconditional Brownian motion W the processes $W' = (W_t - \min_{s\leq t} W_s)_{0\leq t\leq 1}$ and $(|B_t|)_{0\leq t\leq 1}$ are equal in distribution, where B denotes another Brownian motion. Conditioning W to take its minimum at time 1 is equivalent to conditioning W_1' to take the value 0. For B this means that we pass over to a Brownian bridge.

Theorem 1.2. *Under Assumption A, it holds that as $n \to \infty$,*

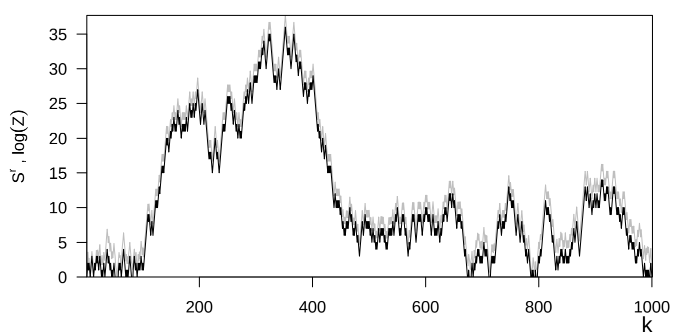
$$\Big(\Big(\frac{1}{\sigma\sqrt{n}} \log Z_{nt}\Big)_{0\le t\le 1} \;\Big|\; Z_n > 0\Big) \xrightarrow{d} |B|$$

in the Skorohod space, where $B = (B_t)_{0\le t\le 1}$ denotes a standard Brownian bridge.

For the linear fractional case convergence of finite dimensional distributions was obtained by Afanasyev [2] (without identifying the limit). This theorem is proved in Section 3. We will now illustrate it by some simulations. The next figure shows two paths corresponding to the path $S_0^c, \dots, S_n^c$ of Figure 1: In black the path $S_0^r, \dots, S_n^r$, given by

$$S_k^r = S_k^c - \min_{j\le k} S_j^c \ ,$$

that is the conditional random walk, reflected at its current minimum, and in grey the path $\log Z_0, \dots, \log Z_n$, where $1 = Z_0, \dots, Z_n$ denotes the branching process, given the environment determined by $S_0^c, \dots, S_n^c$ and conditioned to survive within this environment. The fit of both paths is clearly visible.



The next theorem focuses on the difference between the grey and black processes in the last figure as well as on the magnitude of the population in ladder points of the random walk. It is taken from [4][Theorem 1.4]. Recall that τ_{nt} is the moment, when $S_0, \dots, S_{nt}$ takes its minimum.

Theorem 1.3. *Let* $0 < t_1 < \cdots < t_r < 1$. *For* $i = 1, \ldots, r$ *let*

$$\mu(i) = \min \left\{ j \le i : \inf_{t \le t_j} W_t^c = \inf_{t \le t_i} W_t^c \right\} .$$

Then under Assumption A there are i.i.d. random variables $U_1, \ldots, U_r$ *with values in* $\mathbb{N}$ *and independent of* W^c *such that*

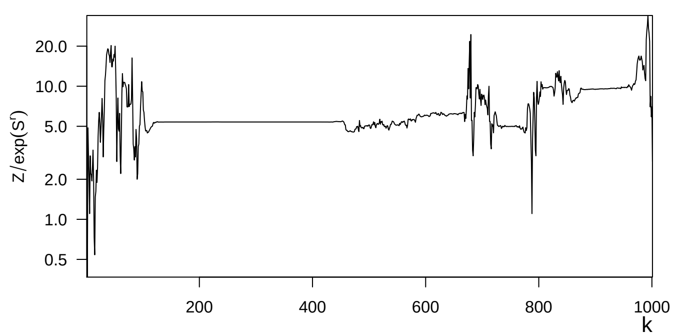
$$\left((Z_{\tau_{nt_1}}, \ldots, Z_{\tau_{nt_r}}) \mid Z_n > 0 \right) \xrightarrow{d} (U_{\mu(1)}, \ldots, U_{\mu(r)})$$

as $n \to \infty$. *Also there are i.i.d. strictly positive random variables* $V_1, \ldots, V_r$ *independent of* W^c *such that*

$$\left(\Big(\frac{Z_{nt_1}}{e^{S_{nt_1} - S_{\tau_{nt_1}}}}, \ldots, \frac{Z_{nt_r}}{e^{S_{nt_r} - S_{\tau_{nt_r}}}} \Big) \mid Z_n > 0 \right) \xrightarrow{d} (V_{\mu(1)}, \ldots, V_{\mu(r)})$$

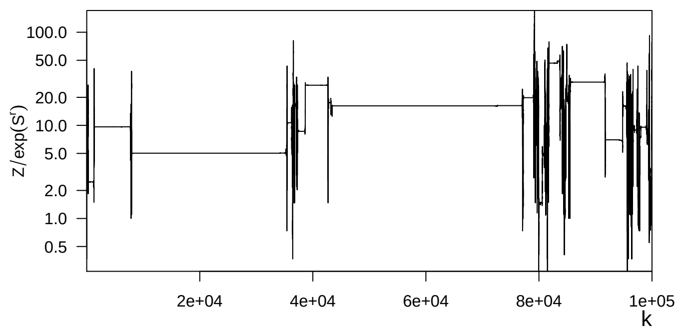
as $n \to \infty$.

The first part confirms that the population size is of order $O(1)$ in ladder points. The meaning of the second statement has already been explained in [4]. Shortly speaking: Within upward excursions $Z_k / \exp(S_k - \min_{j \le k} S_j)$ takes asymptotically a constant value, whereas this value changes in an independent manner from one excursion to the next. This is expressed in the next two pictures. The first shows $Z_k / \exp(S_k^r)$ for the random path $S_0^c, \ldots, S_n^c$ from Figure 1.

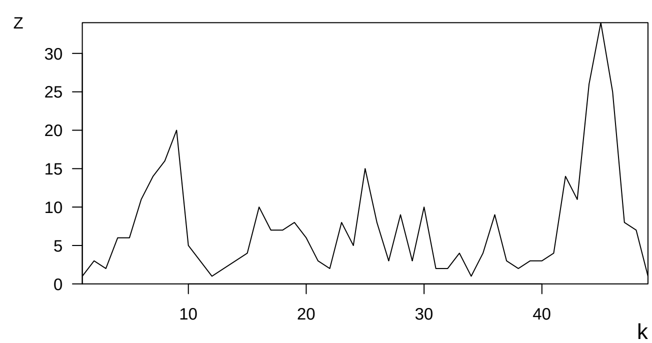


Note that Theorem 1.3 only deals with finite dimensional distributions, it cannot be extended to a functional limit theorem in Skorohod space in a standard manner, since the limiting process would consist of paths being constant within Brownian excursions but independent between different excursions. This

leads to paths which are not càdlàg. This becomes manifest in the next picture with $n = 100000$. Note the heavy oscillations of the path between its constant sections.



The final picture shows the magnitude of the population, when restricted just to the strictly decreasing ladder points of $S_0^c, \ldots, S_n^c$, which are 49 in our $n = 1000$ example. One observes that this process is not just white noise. The dependence structure does not seem to be easily captured.



The rest of the paper is organized as follows. In Section 2 we assemble the facts, relevant for the simulations. In Section 3 we give the proof of Theorem 1.2 in a more general setting.

2 Simulation of branching processes in random environment

In this section, we derive and describe the simulation algorithm which was used to sample the simulations presented in the previous section.

2.1 Simulation of a conditional random walk

Theorem 1.1 says that asymptotically as $n \to \infty$, the associated random walk is distributed like a random walk conditioned on having its minimum at the end. Thus here it is our concern to sample paths $S_0^c, S_1^c, \ldots, S_n^c$. Let us introduce for $n \geq 1$

$$M_n \ = \ \max(S_1, \ldots, S_n) \ .$$

Introducing the dual random walk

$$\hat{S}_{n-i} = S_n - S_i \ , \ 0 \leq i \leq n \ ,$$

we get that

$$\begin{aligned}&\big((S_0, \ldots, S_i, \ldots, S_n) \mid S_n < \min(S_0, \ldots, S_{n-1})\big) \\ &\qquad \stackrel{d}{=} \big(\hat{S}_0, \ldots, \hat{S}_n - \hat{S}_{n-i}, \ldots, \hat{S}_n) \mid \max(\hat{S}_1, \ldots, \hat{S}_n) < 0\big) \ .\end{aligned}$$

Thus a random walk path, conditioned on $\{S_n < \min(S_0, \ldots, S_{n-1})\}$ can be sampled by simulating a random walk path, conditioned on $\{M_n < 0\}$ and then deriving therefrom the dual path. Here, we only consider a simple random walk, i.e.

$$\mathbf{P}(X = 1) = \mathbf{P}(X = -1) = \frac{1}{2} \ .$$

By Markov property, the distribution of X_k conditioned on $\{M_n < 0\}$ is for $1 \leq k \leq n$ given by

$$\begin{aligned}\mathbf{P}(X_k = 1 \mid M_n < 0, S_{k-1} = x) &= \frac{1}{2} \frac{\mathbf{P}(M_n < 0 \mid S_k = x + 1)}{\mathbf{P}(M_n < 0 \mid S_{k-1} = x)} \\ &= \frac{1}{2} \frac{\mathbf{P}(M_{n-k} < -x - 1)}{\mathbf{P}(M_{n-k+1} < -x)} \ . \end{aligned} \tag{2.4}$$

By the reflection principle, the distribution of the maximum is given by

$$\mathbf{P}(M_n \geq y) = \mathbf{P}(S_n \geq y) + \mathbf{P}(S_n > y) \ .$$

For a simple random walk

$$\mathbf{P}(S_n = y) = \mathbf{P}\big(Y = (n-y)/2\big) \ ,$$

where Y is binomially distributed with parameters $(n, 1/2)$. Thus the probability in (2.4) is easily calculated and sampling paths of the conditional random walk is straightforward.

2.2 Geiger construction

In this section, we treat branching processes in varying environment, i.e. the environment $\pi = (q_1, q_2, \ldots)$ is fixed. We denote the underlying probability measure by $\mathbf{P}_\pi(\cdot) = \mathbb{P}(\cdot \mid \Pi = \pi)$.

The Geiger construction allows to construct a branching process in varying environment, conditioned on $\{Z_n > 0\}$, along its ancestral line (see e.g. [11, 4]). Following the notation in [4], we denote by $\mathcal{T}_n$ the set of all ordered rooted trees of height exactly n, $n \in \mathbb{N}_0$ and

$$\mathcal{T}_{\geq n} = \mathcal{T}_n \cup \mathcal{T}_{n+1} \cup \cdots \cup \mathcal{T}_\infty$$

the set of all trees of height at least n.

Let us introduce the operation $[\ ] : \ \mathcal{T}_{\geq n} \to \mathcal{T}_n$ of pruning a tree of height $\geq n$ to a tree of height exactly n by eliminating all nodes of larger height.

A tree with a stem (following [4] called a *trest*) is defined by a pair

$$\mathsf{t} = (t, k_0 k_1 \ldots k_n) \ ,$$

where $t \in \mathcal{T}_{\geq n}$ and $k_0, \ldots, k_n$ are nodes in t such that k_0 is the root (founding ancestor) and k_i is an offspring of k_{i-1}.

The operation

$$\langle t \rangle_n = ([t]_n, k_0(t) \ldots k_n(t))$$

maps an ordered and rooted tree of height at least n to the associated, unique trest of height n, where $k_0(t) \ldots k_n(t)$ is the *leftmost* stem, which can be fitted into $[t]_n$.

Now, we are able to construct the conditional branching tree along its ancestral line. Let

$$\mathsf{T}_{n,\pi} = (T_n, K_0 \ldots K_n)$$

denote a random trest of height n and let for $i = 1, \dots, n$

$$
\begin{aligned}
T_i' &= \text{subtree within } T_n \text{ right to the stem with root } K_{i-1} \\
T_i'' &= \text{subtree within } T_n \text{ left to the stem with root } K_{i-1} \\
R_i &= \text{size of the first generation of } T_i' \\
L_i &= \text{size of the first generation of } T_i''
\end{aligned}
$$

Definition 2.1 (Geiger tree)**.** *A random trest $\mathsf{T}_{n,\pi}$ is called a Geiger tree in the environment π and its distribution is uniquely determined (up to possible offspring of K_n) if the following properties are fulfilled:*

- *The joint distribution of R_i and L_i is given by*

$$
\begin{aligned}
&\mathbf{P}_\pi(R_i = r, L_i = l) \\
&\quad = q_i(r + l + 1) \frac{\mathbf{P}_\pi(Z_n > 0 \mid Z_i = 1)\mathbf{P}_\pi(Z_n = 0 \mid Z_i = 1)^l}{\mathbf{P}_\pi(Z_n > 0 \mid Z_{i-1} = 1)}
\end{aligned}
\tag{2.5}
$$

- *T_i' decomposed at its first generation consists of R_i subtrees τ_{ij}', $j = 1, \dots, R_i$, which are branching trees within the fixed environment $(q_{i+1}, q_{i+2}, \dots)$.*
- *T_i'' consists of L_i subtrees τ_{ij}'', which are branching trees within the fixed environment $(q_{i+1}, q_{i+2}, \dots)$ conditioned on extinction before generation $n - i$.*
- *All pairs (R_i, L_i) and all subtrees τ_{ij}', τ_{ij}'' are independent.*

The following theorem (see [11, 4]) assures that the Geiger and the (pruned) branching tree T conditioned on $\{Z_n > 0\}$ have the same distribution:

Theorem 2.2. *For allmost all π the conditional distribution of $\langle T \rangle_n$ given $\Pi = \pi, Z_n > 0$ is equal to the distribution of $\langle \mathsf{T}_{n,\pi} \rangle_n$.*

2.3 Branching processes with geometric offspring distributions

Following the previous section, we can sample a conditional branching tree in the random environment π using the Geiger construction. Clearly, we require

$$
\rho_{i,n} = \mathbf{P}_\pi(Z_n > 0 \mid Z_i = 1) \ .
$$

In the case of geometric (or more generally linear fractional) offspring distributions, i.e. $q_i(k) = p_i(1-p_i)^k$, $p_i \in [0,1]$, a direct calculation is feasible. Using the formula for the generating function in [15][Equation (6) resp. (24)] yields

$$\rho_{i,n} = \Big(\sum_{k=j}^{n} e^{-(S_j - S_i)}\Big)^{-1} .$$

Then the distribution in (2.5) can be written as

$$\mathbf{P}_\pi(R_i = r, L_i = l) = p_i(1-p_i)^{r+l+1}\frac{\rho_{i,n}(1-\rho_{i,n})^l}{\rho_{i-1,n}} . \tag{2.6}$$

The offspring distribution in generation $1 \le i \le n$ in a subtree conditioned on extinction in generation n is given by

$$\tilde{q}_i(k) = \frac{q_i(k)(1-\rho_{i,n})^k}{1-\rho_{i-1,n}} \ , \ k = 1, 2, \ldots$$

If q is geometric with parameter p_i, then

$$\tilde{q}_i(k) = \frac{p_i(1-p_i)^k(1-\rho_{i,n})^k}{1-\rho_{i-1,n}} = \frac{p_i}{1-\rho_{i-1,n}}\big((1-p_i)(1-\rho_{i,n})\big)^k$$

is again geometric with parameter $1-(1-p_i)(1-\rho_{i,n})$. Note that from the definition of X_i and p_i,

$$e^{X_i} = \frac{1}{p_i} - 1 .$$

Using the Geiger construction, all offspring numbers of indivuals in the conditioned and unconditioned subtrees are independent. As it is well-known, the sum of independent geometrically distributed random variables is negative binomially distributed.

Now, let $Z_i^{(u)}$ be the total number of individuals in the unconditioned trees and $Z_i^{(c)}$ be the total number of individuals in the conditioned trees at generation i. Thus, following Theorem 2.2, we have the following simulation algorithm for a branching process in the varying environment $\pi = (q_1, q_2, \ldots)$, conditioned on $\{Z_n > 0\}$:

- Calculate $\rho_{i,n}$ for $0 \le i \le n$.
- Given $Z_{i-1}^{(u)}$, simulate $Z_i^{(u)}$ as one negative binomially random variable of parameters $(Z_{i-1}^{(u)}, p_{i-1})$.
- Given $Z_{i-1}^{(c)}$, simulate $Z_i^{(c)}$ as one negative binomially random variable of size parameter $Z_{i-1}^{(c)}$ and success probability $1-(1-p_i)(1-\rho_{i,n})$.

- Randomly simulate a pair (R_i, L_i) according to the distribution (2.6). Then add $L_i - 1$-many individuals to $Z_i^{(c)}$ and $R_i - L_i$-many individuals to $Z_i^{(u)}$.
- The total number of individuals in generation i is given by $1 + Z_i^{(u)} + Z_i^{(c)}$.

The simulation amounts to the simulation of one random pair and two independent negative binomially distributed random variables. This allows for very fast simulations.

3 The functional limit theorem

Now we get down to the announced functional limit theorem. The assumptions are the same as in [4]:

Assumption A1. *Let* $\mathbf{E}[X] = 0$.

Assumption A2. *The distribution of X has finite variance with respect to* $\mathbf{P}$ *or (more generally) belongs to the domain of attraction of some zero mean stable law with index $\alpha \in (1, 2]$. It is non-lattice.*

Since $\mathbf{E}[X] = 0$ this implies that there is an increasing sequence of positive numbers

$$a_n \;=\; n^{1/\alpha} \ell_n$$

with a slowly varying sequence $\ell_1, \ell_2, \dots$ such that

$$\frac{S_n}{a_n} \xrightarrow{d} L_1$$

for $n \to \infty$, where L_1 denotes a random variable with the above stable distribution. Note that due to the change of measure X^- always has finite variance and an infinite variance may only arise from X^+. If $\alpha < 2$ this is called the spectrally positive case.

Assumption A3. *For some $\varepsilon > 0$ and some $a \in \mathbb{N}$*

$$\mathbf{E}[(\log^+ \zeta(a))^{\alpha+\varepsilon}] \;<\; \infty \; ,$$

where $\log^+ x = \log(x \vee 1)$.

As is well-known, there is a Lévy process $L = (L_t)_{0\le t\le 1}$ including the stable random variable L_1 above. Let L^c be the corresponding Lévy process conditioned on having its minimum at time $t = 1$. For the precise definition of such a process, we refer to [4]. Again define the process $L^r = (L^r_{0\le t\le 1})$, which is the process L^c reflected at its current minimum and given by

$$L^r_t = L^c_t - \min_{s\le t} L^c_s \ .$$

Theorem 3.1. *Under Assumptions A1 to A3,*

$$\left((\tfrac{1}{a_n} \log Z_{nt})_{0\le t\le 1} \;\middle|\; Z_n > 0 \right) \quad \xrightarrow{d} \quad L^r$$

in the Skorohod space.

The convergence of the finite-dimensional distributions follows directly from known results: From [4][Theorem 1.4] for all $0 \le t \le 1$

$$\left(\frac{1}{a_n} (\log Z_{nt} - S_{nt} - \min_{s\le t} S_{ns}) \;\middle|\; Z_n > 0 \right) \to 0$$

in probability, and from [4][Theorem 1.3]

$$\left(\frac{1}{a_n} S^n \;\middle|\; Z_n > 0 \right) \xrightarrow{d} L^c \ .$$

For tightness, we require the following lemma. Let

$$\eta_i = \zeta_i(1) = \sum_{y=1}^{\infty} y^2 Q_i(y) \Big/ m(Q_i)^2 \ .$$

Lemma 3.2. *Under Assumptions A1 to A3, for every $\upsilon_n = o(n)$ and $\delta > 0$,*

$$e^{-\delta a_n} \sum_{i=0}^{\upsilon_n - 1} \eta_{i+1} e^{S_{\upsilon_n} - S_i} \to 0$$

in probability with respect to $\mathbf{P}$.

Proof. By duality

$$\sum_{i=0}^{\upsilon_n - 1} \eta_{i+1} e^{S_{\upsilon_n} - S_i} \stackrel{d}{=} \sum_{i=1}^{\upsilon_n} \eta_i e^{S_i} \ .$$

Note that $\eta_i \le a^2 e^{-2X_i} + \zeta_i(a)$, where a is the constant from Assumption A3. Recall $M_n = \max(S_1, \ldots, S_n)$. Thus because of $S_i - 2X_i = S_{i-1} - X_i$

$$\begin{aligned}\sum_{i=1}^{\upsilon_n} \eta_i e^{S_i} &\le a^2 \sum_{i=1}^{\upsilon_n} e^{S_i - 2X_i} + \sum_{i=1}^{\upsilon_n} \zeta_i(a) e^{S_i} \\ &\le a^2 e^{M^+_{\upsilon_n}} \sum_{i=1}^{\upsilon_n} e^{-X_i} + e^{M_{\upsilon_n}} \sum_{i=1}^{\upsilon_n} \zeta_i(a) \ .\end{aligned}$$

As $\mathbf{E}[e^{-X}] = \gamma^{-1}\mathbb{E}[e^X e^{-X}] = \gamma^{-1} < \infty$, the law of large numbers implies

$$\sum_{i=1}^{\upsilon_n} e^{-X_i} = O(\upsilon_n) \quad \mathbf{P}\text{-a.s.}$$

Assumption A3 and the Borel-Cantelli lemma imply $(\log^+ \zeta_i(a))^{\alpha+\varepsilon} = O(i)$ respectively

$$\zeta_i(a) = O\Big(\exp\big(i^{1/(\alpha+\varepsilon)}\big)\Big) \quad \mathbf{P}\text{-a.s.}$$

Recall that $\upsilon_n = o(n)$ and $a_n = n^{1/\alpha} l(n)$, where $l(n)$ is slowly varying. Thus for every $\delta > 0$,

$$\sum_{i=1}^{\upsilon_n} \zeta_i(a) = O\Big(\upsilon_n \exp\big(n^{1/(\alpha+\varepsilon)}\big)\Big) = O\big(e^{\delta a_n/3}\big) \quad \mathbf{P}\text{-a.s.}$$

As a consequence of the functional limit theorem [14][Theorem 16.14], $\frac{1}{a_n} M_n$ converges in distribution with respect to $\mathbf{P}$, thus

$$M_{\upsilon_n} = O_P\big(a_{\upsilon_n}\big) = o_P\big(a_n\big) \ .$$

Using these results

$$\sum_{i=1}^{\upsilon_n} \eta_i e^{S_i} = O_P\big(e^{\delta a_n/2}\big) \ .$$

This yields the claim. □

The following lemma immediately results.

Lemma 3.3. *Under Assumptions A1 to A3, for every $\upsilon_n = o(n)$ and $\delta > 0$,*

$$e^{\delta a_n} \mathbf{P}(Z_{\upsilon_n} > 0 \mid \Pi) \to \infty$$

in probability with respect to $\mathbf{P}$.

Proof. We use the standard lower estimate for the survival probability (see e.g. [7])

$$\mathbf{P}(Z_{\upsilon_n} > 0 \mid \Pi) \geq \frac{1}{e^{-S_{\upsilon_n}} + \sum_{k=0}^{\upsilon_n - 1} \eta_{k+1} e^{-S_k}} \qquad \text{a.s.}$$

Thus it remains to show that

$$e^{-\delta a_n}\Big(e^{-S_{\upsilon_n}} + \sum_{k=0}^{\upsilon_n - 1} \eta_{k+1} e^{-S_k}\Big) \to 0$$

which is proved in the same way as Lemma 3.2. □

Proof of Theorem 3.1. It remains to prove tightness. We use Aldous' criterium (see e.g. [14], p. 314) and show that for any sequence of positive constants $\upsilon_n = o(n)$, any sequence of stopping times $\kappa_1, \kappa_2, \ldots$, with $\kappa_n \leq n$ and any $\delta > 0$

$$\mathbb{P}\big(\tfrac{1}{a_n} |\log Z_{(\kappa_n + \upsilon_n)\wedge n} - \log Z_{\kappa_n}| > \delta \mid Z_n > 0\big) \to 0 \tag{3.7}$$

for $n \to \infty$. First let us fix $0 < s < 1$ and additionally assume $\kappa_n \leq sn$. We show that in

$$\begin{aligned} &\mathbb{P}(\tfrac{1}{a_n} |\log Z_{\kappa_n + \upsilon_n} - \log Z_{\kappa_n}| > \delta \mid Z_n > 0) \\ &\quad = \mathbb{P}(Z_{\kappa_n + \upsilon_n} > e^{\delta a_n} Z_{\kappa_n} \mid Z_n > 0) + \mathbb{P}(Z_{\kappa_n + \upsilon_n} < e^{-\delta a_n} Z_{\kappa_n} \mid Z_n > 0) \end{aligned} \tag{3.8}$$

both right-hand terms converge to 0.

Let us treat the first summand in (3.8). Because of the independence properties of a BPRE, it follows that for stopping times κ_n

$$\begin{aligned} &\mathbb{P}(Z_{\kappa_n + \upsilon_n} > e^{\delta a_n} Z_{\kappa_n}, Z_n > 0) \\ &\quad = \sum_{z \geq 1, k \leq sn} \mathbb{P}(\kappa_n = k, Z_k = z) \mathbb{P}_z(Z_{\upsilon_n} > e^{\delta a_n} z, Z_{n-k} > 0) \end{aligned} \tag{3.9}$$

with $\mathbb{P}_z(\cdot) = \mathbb{P}(\cdot \mid Z_0 = z)$. As to the right-hand probability we distinguish two possibilities: Either one of the z initial individuals has at least one offspring in generation $n-k$ and more than $e^{\delta a_n}$ offspring in generation υ_n. Or it has at least one offspring in generation $n-k$ and the other $z-1$ individuals have more than $(z-1)e^{\delta a_n}$ offspring in generation υ_n. Thus a.s.

$$\begin{aligned} \mathbb{P}_z(Z_{\upsilon_n} > e^{\delta a_n} z, Z_{n-k} > 0 \mid \Pi) &\leq z\mathbb{P}(Z_{\upsilon_n} > e^{\delta a_n}, Z_{n-k} > 0 \mid \Pi) \\ &\quad + z\mathbb{P}(Z_{n-k} > 0 \mid \Pi) \cdot \mathbb{P}_{z-1}(Z_{\upsilon_n} > (z-1)e^{\delta a_n} \mid \Pi)\ . \end{aligned} \tag{3.10}$$

As to the last term by means of Markov's inequality a.s.

$$\mathbb{P}_{z-1}(Z_{\upsilon_n} > (z-1)e^{\delta a_n} \mid \Pi) \leq \frac{1}{(z-1)e^{\delta a_n}} \mathbb{E}_{z-1}[Z_{\upsilon_n} \mid \Pi] = e^{-\delta a_n} e^{S_{\upsilon_n}} \ ,$$

thus

$$\begin{aligned}
&\mathbb{E}\big[\mathbb{P}(Z_{n-k} > 0 \mid \Pi) \cdot \mathbb{P}_{z-1}(Z_{\upsilon_n} > (z-1)e^{\delta a_n} \mid \Pi)\big] \\
&\quad \leq \mathbb{E}\big[Z_{\upsilon_n} \mathbb{P}(Z_{n-k} > 0 \mid \Pi, Z_{\upsilon_n} = 1) \cdot 1 \wedge (e^{-\delta a_n} e^{S_{\upsilon_n}}) \mid \Pi)\big] \\
&\quad = \mathbb{E}\big[e^{S_{\upsilon_n}} \cdot 1 \wedge e^{S_{\upsilon_n} - \delta a_n}\big] \mathbb{E}[\mathbb{P}(Z_{n-k} > 0 \mid \Pi, Z_{\upsilon_n} = 1)] \\
&\quad = \gamma^{\upsilon_n} \mathbf{E}\big[1 \wedge e^{S_{\upsilon_n} - \delta a_n}\big] \mathbb{P}(Z_{n-k-\upsilon_n} > 0) \ .
\end{aligned}$$

As $S_{\upsilon_n} - \delta a_n \to -\infty$ in probability with respect to $\mathbf{P}$ it follows by dominated convergence for $n \to \infty$

$$\mathbf{E}\big[\big(1 \wedge e^{S_{\upsilon_n} - \delta a_n}\big)\big] \to 0 \ .$$

Also, applying the remarks after [4][Theorem 1.1], $\mathbb{P}(Z_n > 0) \sim \theta \gamma^n / b_n$, where b_n is regularly varying with exponent $1 - \alpha^{-1}$. Thus

$$\gamma^{\upsilon_n} \mathbb{P}(Z_{n-k-\upsilon_n} > 0) = O\big(\mathbb{P}(Z_{n-k} > 0)\big)$$

uniformly in $k \leq sn$ and consequently

$$\mathbb{E}\big[\mathbb{P}(Z_{n-k} > 0 \mid \Pi) \cdot \mathbb{P}_{z-1}(Z_{\upsilon_n} > (z-1)e^{\delta a_n} \mid \Pi)\big] = o(\mathbb{P}(Z_{n-k} > 0))$$

uniformly in $z \geq 1$ and $k \leq sn$.

Next, let us show the same statement for the other term in (3.10). For this, we will use [4][Theorem 4.2]. Let $\tilde{\mathsf{T}}$ be the LPP-trest defined in [16, 4] and $\tilde{Z}_n$ its population size in generation n. As above, let

$$\tau_n = \min\{k \leq n : S_k = \min(S_0, \ldots, S_n)\} \ .$$

Let $m \in \mathbb{N}$ be fixed. Without loss of generality, we write n instead of $n - k \geq (1-s)n$ in the following estimates.

In [4][Equation (4.9)], it is shown that

$$\mathbf{E}\big[\tilde{Z}_n \mid \Pi\big] = 1 + \sum_{i=0}^{n-1} \eta_{i+1} e^{S_n - S_i} \ .$$

Using this together with Markov inequality yields

$$\begin{aligned}
&\mathbf{P}\big(\tilde{Z}_{\upsilon_n} > e^{\delta a_n}, \tau_{n-m} = n-m\big) \le \mathbf{E}\Big[1 \wedge \Big(e^{-\delta a_n}\mathbf{E}\big[\tilde{Z}_n \mid \Pi\big]\Big); \tau_{n-m} = n-m\Big] \\
&\le \mathbf{E}\Big[1 \wedge \Big(e^{-\delta a_n}\mathbf{E}\big[\tilde{Z}_n \mid \Pi\big]\Big); S_{n-m} < S_{\upsilon_n}, S_{\upsilon_n+1}, \dots, S_{n-m-1}\Big] \\
&= \mathbf{E}\Big[1 \wedge \Big(e^{-\delta a_n}\big(1 + \sum_{i=0}^{\upsilon_n-1} \eta_{i+1} e^{S_{\upsilon_n}-S_i}\big)\Big)\Big]\mathbf{P}(\tau_{n-m-\upsilon_n} = n-m-\upsilon_n)\ .
\end{aligned}$$

By [4][Lemma 2.2] $\mathbf{P}(\tau_n = n)$ is regularly varying, thus, as $n \to \infty$, $\mathbf{P}(\tau_{n-m-\upsilon_n} = n-m-\upsilon_n) \sim \mathbf{P}(\tau_n = n)$. From Lemma 3.2 and the dominated convergence theorem, it results that

$$\mathbf{E}\Big[1 \wedge \Big(e^{-\delta a_n}\big(1 + \sum_{i=0}^{\upsilon_n-1} \eta_{i+1} e^{S_{\upsilon_n}-S_i}\big)\Big)\Big] \to 0\ .$$

Altogether, as $n \to \infty$,

$$\mathbf{P}(\tilde{Z}_{\upsilon_n} > e^{\delta a_n} \mid \tau_{n-m} = n-m) \to 0\ .$$

Recall that by definition of υ_n, $n - \upsilon_n \to \infty$. Thus using [4][Theorem 4.2] we get that uniformly in $k \le sn$

$$\mathbb{P}(Z_{\upsilon_n} > e^{\delta a_n}, Z_{n-k} > 0) = o\big(\mathbb{P}(Z_{n-k} > 0)\big)\ .$$

Altogether, from (3.10)

$$\mathbb{P}_z(Z_{\upsilon_n} > e^{\delta a_n} z, Z_{n-k} > 0) = o\big(z\mathbb{P}(Z_{n-k} > 0)\big)$$

uniformly in $z \ge 1$ and $k \le sn$. Applying this result to (3.9) and changing to the measure $\mathbf{P}$ yields for any $\varepsilon > 0$

$$\begin{aligned}
&\mathbb{P}(Z_{\kappa_n+\upsilon_n} > e^{\delta a_n} Z_{\kappa_n}, Z_n > 0) \le \varepsilon \sum_{z\ge 1, k\le sn} \mathbb{P}(\kappa_n = k, Z_k = z) z \mathbb{P}(Z_{n-k} > 0) \\
&\le \varepsilon \sum_{k \le sn} \gamma^k \mathbf{E}\big[e^{-S_k} Z_k; \kappa_n = k\big]\mathbb{P}(Z_{n-k} > 0) \\
&= \varepsilon \sum_{k \le sn} \gamma^k \mathbf{E}\big[e^{-S_k} \mathbb{E}[Z_k \mid \Pi]; \kappa_n = k\big]\mathbb{P}(Z_{n-k} > 0) \\
&= \varepsilon \sum_{k \le sn} \gamma^k \mathbf{P}(\kappa_n = k)\mathbb{P}(Z_{n-k} > 0)\ ,
\end{aligned} \tag{3.11}$$

if n is large enough. Finally, applying [4][Theorem 1.1], $\mathbb{P}(Z_n > 0) \sim \theta\gamma^n/b_n$ and thus for large n

$$\mathbb{P}(Z_{\kappa_n+\upsilon_n} > e^{\delta a_n} Z_{\kappa_n} \mid Z_n > 0) \le 2\varepsilon \frac{b_n}{b_{(1-s)n}}\ .$$

Taking the limit $n \to \infty$ and then $\varepsilon \to 0$

$$\mathbb{P}(Z_{\kappa_n+\upsilon_n} > e^{\delta a_n} Z_{\kappa_n} \mid Z_n > 0) \to 0 \ .$$

Next, let us turn to the second summand in (3.8). Applying similar arguments as before,

$$\begin{aligned}\mathbb{P}_z(Z_{\upsilon_n} < ze^{-\delta a_n}, Z_{n-k} > 0 \mid \Pi) &\leq z\mathbb{P}(Z_{\upsilon_n} < e^{-\delta a_n}, Z_{n-k} > 0 \mid \Pi) \\ &+ z\mathbb{P}(Z_{n-k} > 0 \mid \Pi)\mathbb{P}_{z-1}\big(Z_{\upsilon_n} < (z-1)e^{-\delta a_n} \mid \Pi\big) \text{ a.s.}\end{aligned}$$

As $e^{-\delta a_n} < 1$, the first right-hand term vanishes, also for $ze^{-\delta a_n} < 1$ the left-hand side is 0. Thus the inequality becomes

$$\begin{aligned}&\mathbb{P}_z(Z_{\upsilon_n} < ze^{-\delta a_n}, Z_{n-k} > 0 \mid \Pi) \\ &\quad\leq z\mathbf{1}_{\{z \geq e^{\delta a_n}\}}\mathbb{P}(Z_{n-k} > 0 \mid \Pi)\mathbb{P}_{z-1}\big(Z_{\upsilon_n} < (z-1)e^{-\delta a_n} \mid \Pi\big) \text{ a.s.}\end{aligned}$$

Conditioned on the environment, all $(z-1)$-many subtrees of the branching process are independent. Of these subtrees, at most $(z-1)e^{-\delta a_n}$-many may survive until generation υ_n. Thus, letting Y be a binomially distributed random variable with parameters $(z-1,p)$, $p = \mathbb{P}(Z_{\upsilon_n} > 0 \mid \Pi)$, by means of Chebyshev's inequality for $z > 1$

$$\begin{aligned}\mathbb{P}_{z-1}\big(Z_{\upsilon_n} < (z-1)e^{-\delta a_n} \mid \Pi\big)\mathbf{1}_{\{z \geq e^{\delta a_n}\}} &\leq \mathbb{P}\big(Y < (z-1)e^{-\delta a_n} \mid \Pi\big)\mathbf{1}_{\{z \geq e^{\delta a_n}\}} \\ &\leq \mathbf{1}_{\{e^{-\delta a_n} > p/2\}} + \mathbb{P}\big(Y < (z-1)p/2 \mid \Pi\big)\mathbf{1}_{\{z \geq e^{\delta a_n}\}} \\ &\leq \mathbf{1}_{\{e^{\delta a_n}p < 2\}} + \frac{4}{(z-1)p}\mathbf{1}_{\{z \geq e^{\delta a_n}\}} \\ &\leq \frac{2}{e^{\delta a_n}p} + \frac{4}{(e^{\delta a_n}-1)p} \leq \frac{6}{(e^{\delta a_n}-1)p} \ .\end{aligned}$$

It follows

$$\begin{aligned}\mathbb{P}_z(Z_{\upsilon_n} < ze^{-\delta a_n}, Z_{n-k} > 0) &\leq 6z\mathbb{E}\big[\mathbb{P}(Z_{n-k} > 0 \mid \Pi)\big(1 \wedge ((e^{\delta a_n}-1)p)^{-1}\big)\big] \\ &\leq 6z\mathbb{E}\big[Z_{\upsilon_n}\mathbb{P}(Z_{n-k-\upsilon_n} > 0 \mid \Pi, Z_{\upsilon_n} = 1)\big(1 \wedge ((e^{\delta a_n}-1)p)^{-1}\big)\big] \\ &= 6z\mathbb{E}\big[e^{S_{\upsilon_n}}\big(1 \wedge ((e^{\delta a_n}-1)p)^{-1}\big)\big]\mathbb{P}(Z_{n-k-\upsilon_n} > 0) \\ &= 6z\gamma^{\upsilon_n}\mathbf{E}\big[\big(1 \wedge ((e^{\delta a_n}-1)p)^{-1}\big)\big]\mathbb{P}(Z_{n-k-\upsilon_n} > 0) \ .\end{aligned}$$

From Lemma 3.3 and dominated convergence, we obtain

$$\mathbf{E}\big[\big(1 \wedge ((e^{\delta a_n}-1)p)^{-1}\big)\big] \to 0$$

for $n \to \infty$, thus as above

$$\mathbb{P}_z(Z_{\upsilon_n} < e^{-\delta a_n} z, Z_{n-k} > 0) = o\big(z\mathbb{P}(Z_{n-k} > 0)\big)$$

uniformly in $z \geq 1$ and $k \leq sn$, and in much the same way as above we may conclude

$$\mathbb{P}(Z_{\kappa_n + \upsilon_n} < e^{-\delta a_n} Z_{\kappa_n} \mid Z_n > 0) \to 0$$

as $n \to \infty$. From (3.8) we see that (3.7) is satisfied for any sequence of stopping times κ_n such that $\kappa_n \leq sn$ with some $s < 1$.

Finally, we show that (3.7) also holds for all stopping times $\kappa_n \leq n$. Let $0 \leq s < 1$. Then we get that for large n

$$\begin{aligned}
&\mathbb{P}\big(\tfrac{1}{a_n}\big|\log Z_{(\kappa_n+\upsilon_n)\wedge n} - \log Z_{\kappa_n}\big| > \delta \mid Z_n > 0\big) \\
&\quad \leq \mathbb{P}\big(\tfrac{1}{a_n}\big|\log Z_{\kappa_n \wedge sn + \upsilon_n} - \log Z_{\kappa_n \wedge sn}\big| > \delta \mid Z_n > 0\big) \\
&\qquad + 2\mathbb{P}\Big(\tfrac{1}{a_n}\sup_{s \leq t \leq 1}\big|\log Z_{nt} - \log Z_n\big| > \delta \mid Z_n > 0\Big) \ .
\end{aligned}$$

Since $\kappa_n \wedge sn$ is again a stopping time, taking the lim sup, the first term above vanishes for every $0 \leq s < 1$. Thus it is enough to show that the second term can be made arbitrarily small in the limit, if we choose s close enough to 1. Now

$$\begin{aligned}
&\mathbf{1}_{\{Z_n > 0\}} \tfrac{1}{a_n} \sup_{s \leq t \leq 1} \big|\log Z_{nt} - \log Z_n\big| \\
&\qquad \leq \mathbf{1}_{\{Z_n > 0\}} \tfrac{1}{a_n} \sup_{s \leq t \leq 1} \big(\log Z_{nt} + \log Z_n\big) \\
&\qquad \leq \tfrac{1}{a_n} \sup_{0 \leq t \leq 1} \big(\log^+ Z_{nt} - (S_{nt} - S_n)\big) + \tfrac{1}{a_n} \sup_{s \leq t \leq 1} |S_{nt} - S_n| + \tfrac{1}{a_n} \log^+ Z_n \ .
\end{aligned}$$

Conditioned on $Z_n > 0$, Z_n has a limiting distribution on $\mathbb{N}$, see [4][Theorem 1.2]. Thus, as $n \to \infty$, for $\delta > 0$

$$\lim_{n \to \infty} \mathbb{P}\big(\tfrac{1}{a_n} \log^+ Z_n > \delta \mid Z_n > 0\big) = 0 \ .$$

As to the second term, using [4][Theorem 1.3] and letting $n \to \infty$, for every $\varepsilon > 0$,

$$\mathbb{P}\big(\tfrac{1}{a_n} \sup_{s \leq t \leq 1} |S_n - S_{nt}| > \delta \mid Z_n > 0\big) < \varepsilon$$

if only s is chosen close enough to 1.

Finally, as conditioned on Π, $Z_{nt}e^{-S_{nt}}$ is a non-negative martingale with mean 1, we use the Doob inequality to get that

$$\begin{aligned}
&\mathbb{P}\Big(\tfrac{1}{a_n}\sup_{0\le t\le 1}\big(\log Z_{nt}-(S_{nt}-S_n)\big)>\delta, Z_n>0\Big)\\
&\quad\le \mathbb{P}\Big(\sup_{0\le t\le 1} Z_{nt}e^{-S_{nt}+S_n}>e^{\delta a_n}\Big)\\
&\quad= \mathbb{E}\Big[\mathbb{P}\Big(\sup_{0\le t\le 1} Z_{nt}e^{-S_{nt}}>e^{\delta a_n-S_n}\mid \Pi\Big)\Big]\\
&\quad\le \mathbb{E}\big[e^{S_n-\delta a_n}\big]=\gamma^n e^{-\delta a_n}\ .
\end{aligned}$$

Again by [4][Theorem 1.1], $\gamma^n e^{-\delta a_n}=o\big(\mathbb{P}(Z_n>0)\big)$. Altogether, this yields

$$\limsup_{n\to\infty}\mathbb{P}\Big(\tfrac{1}{a_n}\sup_{s\le t\le 1}\big|\log Z_{nt}-\log Z_n\big|>3\delta\mid Z_n>0\Big)\le\varepsilon$$

if only s is close enough to 1. Thus we have proved (3.7) for every sequence of stopping times $\kappa_n\le n$ which yields tightness. □

References

[1] Afanasyev, V. I. (1980). Limit theorems for a conditional random walk and some applications. *Diss. Cand. Sci.*, Moscow, MSU.

[2] Afanasyev, V. I. (2001). Limit theorems for an intermediately subcritical and a strongly subcritical branching process in a random environment. *Discrete Math. Appl.* **11**, 105-131.

[3] Afanasyev, V. I., Böinghoff, Ch., Kersting, G., and Vatutin, V. A. (2012). Limit theorems for weakly subcritical branching processes in random environment. *J. Theoret. Probab.* **25**, 703–732.

[4] Afanasyev, V. I., Böinghoff, Ch., Kersting, G., and Vatutin, V. A. (2012). Conditional limit theorems for intermediately subcritical branching processes in random environment *http://arxiv.org/abs/1108.2127*

[5] Afanasyev, V. I., Geiger, J., Kersting, G., and Vatutin, V. A. (2005). Criticality for branching processes in random environment. *Ann. Probab.* **33**, 645–673.

[6] Afanasyev, V. I., Geiger, J., Kersting, G., and Vatutin, V. A. (2005). Functional limit theorems for strongly subcritical branching processes in random environment. *Stochastic Process. Appl.* **115**, 1658–1676.

[7] Agresti, A. (1975). On the extinction times of varying and random environment branching processes. *J. Appl. Probab.* **12**, 39–46.

[8] Athreya, K.B. and Karlin, S. (1971). On branching processes with random environments: I, II. *Ann. Math. Stat.* **42**, 1499–1520, 1843–1858.

[9] Birkner, M., Geiger, J., and Kersting, G. (2005). Branching processes in random environment - a view on critical and subcritical cases. Proceedings of the DFG-Schwerpunktprogramm *Interacting Stochastic Systems of High Complexity*, Springer, Berlin, 265–291.

[10] Dekking, F.M. (1988). On the survival probability of a branching process in a finite state i.i.d. environment. *Stochastic Process. Appl.* **27**, 151–157.

[11] Geiger, J. (1999) Elementary new proofs of classical limit theorems for Galton-Watson processes. *J. Appl. Probab.* **36**, 301–309.

[12] Geiger, J., Kersting, G., and Vatutin, V.A. (2003). Limit theorems for subcritical branching processes in random environment. *Ann. Inst. H. Poincaré Probab. Statist.***39**, 593–620.

[13] Guivarc'h, Y. and Liu, Q. (2001). Propriétés asymptotiques des processus de branchement en environnement aléatoire. *C. R. Acad. Sci. Paris Sér. I Math.* **332**, 339–344.

[14] Kallenberg, O.. Foundations of Modern Probability. Springer-Verlag, London (2001).

[15] Kozlov, M.V. (2006). On large deviations of branching processes in a random environment: geometric distribution of descendants. *Discrete Math. Appl.* **16**, 155–174.

[16] Lyons, R., Pemantle, R., and Peres, Y. (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.* **23**, 1125–1138.

[17] Smith, W.L. and Wilkinson, W.E. (1969). On branching processes in random environments. *Ann. Math. Stat.* **40**, 814–827.

[18] Vatutin, V. A. (2004). A limit theorem for an intermediate subcritical branching process in a random environment. *Theory Probab. Appl* **48**, 481-492.